心理科学进展　2020, Vol. 28, No. 7, 1133–1140
Advances in Psychological Science　DOI: 10.3724/SP.J.1042.2020.01133

•研究简报(Research Reports)•

声乐与器乐情绪加工的 ERP 研究

江　俊[1]　张伟霞[2]　王莞琪[3]

([1]上海师范大学音乐学院; [2]上海师范大学教育学院; [3]上海师范大学商学院, 上海　200234)

摘　要　借助跨通道情绪启动范式考察声乐与器乐情绪加工的电生理差异。启动刺激是声乐和器乐曲(小提琴演奏), 目标刺激是与音乐情绪一致或不一致的面孔表情图片。结果显示, 与一致条件相比, 与面孔情绪不一致的器乐曲诱发了 N400, 而与面孔情绪不一致的声乐曲诱发了 LPC。这些结果表明, 声乐与器乐情绪加工的电生理反应存在差异。

关键词　声乐; 器乐; 情绪加工; N400; LPC
分类号　B842; B845

1　引言

聆听音乐是现代人生活中的重要组成部分。声乐(即歌曲)与器乐(即纯音乐)是日常聆听中的2种音乐类型。然而, 有些人喜欢聆听器乐, 有些人则偏爱聆听声乐, 认为歌词可以帮助理解音乐意义, 如音乐的情绪意义。

那么, 听者对声乐和器乐情绪加工是否存在差异? 行为研究表明, 声乐与器乐情绪加工存在差异。在 Stratton 和 Zalanowski (1994)研究中, 大学生聆听一首表现悲伤情绪的歌曲、歌曲的纯音乐或歌词版本, 在 10 点等级(1 = 非常不愉快, 10 = 非常愉快)上对刺激的愉悦度进行主观评定。结果显示, 大学生对纯音乐愉悦度的评定显著高于对歌曲和歌词的评定。通过 5 点等级评定, Brattico 等人(2011)也发现, 被试对纯音乐的愉悦度评定高于对歌曲的评定。但是, 也有矛盾的研究结果。Loui, Bachorik, Li 和 Schlaug (2013)让被试聆听 16 首歌曲及其器乐版本, 对音乐愉悦度进行 5 点等级评定。研究者发现, 听者对声乐与器乐版本的愉悦度评定无明显差异。

以上研究从情绪效价视角探究声乐与器乐情绪加工, 还有一些研究从情绪强度视角探讨这个问题。在 Ali 和 Peynircioğlu (2006)研究中, 大学生聆听 16 首歌曲及其对应的器乐版本。在听完音乐后, 被试评定音乐表现的情绪强度(1 = 一点也不, 9 = 非常)。结果表明, 纯音乐的情绪强度高于歌曲的情绪强度。然而, Franco, Chew 和 Swaine (2017)得出不同的结论。研究者让 3 组大学生分别聆听纯音乐、有纯音乐的歌曲和无纯音乐的歌曲, 评定音乐情绪强度(1 = 一点也不, 7 = 非常)。研究表明, 被试对不同音乐情绪强度的评定基本相同。

已有研究结果的不一致可能由于行为指标比较主观、不敏感引起。那么, 使用相对客观、较为敏感的脑电指标(Kraus & Horowitz-Kraus, 2014; Luck, 2014)有可能更好揭示声乐与器乐情绪加工是否存在差异。跨通道情绪启动范式被广泛运用于情绪加工的 ERP 研究中(比如, Diamond & Zhang, 2016; Steinbeis & Koelsch, 2011)。这类研究主要聚焦 N400 和 LPC 成分。N400 反映了语义或概念的整合加工(Kutas & Federmeier, 2011), 其峰值在 400 ms 左右达到最大, 通常分布在中央和顶区电极位置, 稍微偏右侧化。尽管如此, 在跨通道情绪启动范式研究中, 与启动音乐(Goerlich et al., 2012; Steinbeis & Koelsch, 2011)、言语韵律(Paulmann & Pell, 2010; Schirmer, Kotz, & Friederici, 2005)和面孔图片(Diamond & Zhang, 2016; Föcker & Röder, 2019)情绪不一致的目标刺激诱发了更大的 N400。这表明, N400 也反映了对

收稿日期: 2019-09-23
通信作者: 江俊, E-mail: jjxytc@163.com

情绪冲突的检测与整合(Zhang, Lawson, Guo, & Jiang, 2006; Zhang, Li, Gold, & Jiang, 2010)。LPC是情绪刺激呈现后 300~1000 ms 出现的正波，峰值在 500~600 ms 达到最大，通常分布在中线和顶区电极位置，主要反映了对情绪信息的注意和整合加工(Hajcak, MacNamara, & Olvet, 2010)。具体来说，与中性刺激相比，情绪刺激如图片(Hajcak & Olvet, 2008)和言语韵律(Paulmann & Uskul, 2017)会诱发整个头皮产生一个较大的 LPC。由于情绪刺激具有较强的动机意义，因而引起了更多的注意(Schuppet al., 2000)。同样地，与启动言语韵律(Pell et al., 2015; 郑志伟，黄贤军, 2013)和面孔图片(Diamond & Zhang, 2016)情绪不一致的目标刺激诱发大脑双侧顶枕区或整个头皮产生了更大的LPC。这说明，LPC 与注意分配和情绪整合相关(Herring, Taylor, White, & Crites Jr, 2011; Zhang, Kong, & Jiang, 2012; Zhang et al., 2010)。

最近，张伟霞、王莞琪、周临舒和蒋存梅(2018)通过 ERP 技术考察了歌词对音乐情绪加工的影响。研究表明，在无歌词的哼唱启动下，大脑整个头皮产生了 N400；在带有歌词音乐启动下，大脑整个头皮产生了 LPC。迄今为止，尚未有研究从电生理角度探究声乐与器乐情绪加工的差异。基于此，本研究通过跨通道情绪启动范式对此进行探究。为了排除声乐与器乐曲在音乐形态上的差异，本研究选取相同的音乐，分别由人声演唱和小提琴演奏，由此形成声乐和器乐 2 个版本。既然歌词的出现会诱发 LPC，而无歌词的哼唱诱发出 N400 (张伟霞 等, 2018)，那么我们预期，声乐和器乐情绪加工可能会分别产生 LPC 和 N400。

2 方法

2.1 被试

根据已有研究(张伟霞 等, 2018)中情绪一致性的效应量[$F(1, 15) = 5.17$, $\eta_p^2 = 0.26$, $d = 1.19$]，为了在 0.01 水平上达到 90%的统计检验力，G*Power 3.1.9.4 软件(Faul, Erdfelder, Lang, & Buchner, 2007)非参数配对 t 检验计算出的样本量至少为 17。于是，本实验招募了 20 名未受过专业音乐训练的普通大学生作为被试。由于 3 名被试脑电伪迹过多，在数据分析时被剔除，因此有效被试为 17 人(男生 9 人；年龄 = 24.65 ± 1.31 岁)。被试均报告为右利手，听力和(矫正)视力正常，无脑损伤和精神病史。所有被试签署了知情同意书，并获得一定报酬。

2.2 刺激

刺激包含启动和目标刺激。启动刺激是从声乐教材中挑选的 50 个表达正性情绪和 50 个表达负性情绪的外国歌剧片段(时长 10~25 s，平均 17 s)。歌剧片段分别由一名受过 18 年声乐或小提琴训练的研究生用中文演唱或小提琴演奏并进行录制，由此排除了声乐和器乐在音乐形态上的差异。本实验包含 100 个声乐片段和 100 个器乐片段。这些音乐片段使用单声道采样，采样率为 22050 Hz。最后，使用 Adobe Audition CS6 软件将音量标准化到–7 dB，同时淡出 1 s。

目标刺激是从《中国面孔表情图片系统》(龚栩，黄宇霞，王妍，罗跃嘉, 2011)中挑选的 2 张女性面孔图片，分别表现快乐和悲伤情绪；2 张图片的平均情绪认同度均为 97.98%，其平均情绪强度分别为 7.24 和 5.88。每个音乐与 2 张面孔图片进行配对(见图 1)，因而形成 200 个情绪效价一致和 200 个情绪效价不一致的配对。

为确保刺激的有效性，我们实施了 3 个前测。第一，额外的 16 名普通大学生对 400 个音乐−图片配对的情绪效价一致性(1 = 非常不一致，9 = 非常一致)进行评定。对于每个音乐片段，配对 t 检验被用于检验情绪效价一致与不一致配对的评定是否具有显著差异。240 个配对被作为正式的刺激：不管启动刺激是声乐[$t(59) = 24.73$, $p < 0.001$, $d = 3.19$]还是器乐[$t(59) = 21.56$, $p < 0.001$, $d = 2.78$]，听者对情绪效价一致配对的评定(声乐: $M = 7.21$, $SD = 0.79$；器乐: $M = 6.94$, $SD = 0.87$)均高于对情绪效价不一致配对的评定(声乐：$M = 2.72$, $SD = 0.80$；器乐: $M = 3.15$, $SD = 0.77$)。以情绪效价一致性(情绪效价一致、情绪效价不一致)和音乐类型(声乐、器乐)为被试内变量进行的方差分析结果显示，情绪效价一致性的主效应[$F(1, 59) = 822.81$, $p < 0.001$, $\eta_p^2 = 0.93$]以及情绪效价一致性和音乐类型的交互效应[$F(1, 59) = 11.06$, $p = 0.002$, $\eta_p^2 = 0.16$]显著，音乐类型的主效应不显著，$F(1, 59) = 1.14$, $p = 0.290$。这些结果表明实验材料是有效的，同时，2 种音乐类型的任务难度相似。此外，这些大学生对声乐中歌词清晰度(1 = 不清楚，5 = 清楚)的平均评定高于 4 ($M = 4.45$, $SD = 0.23$)。第二，额外的 16 名音乐专业学生(均

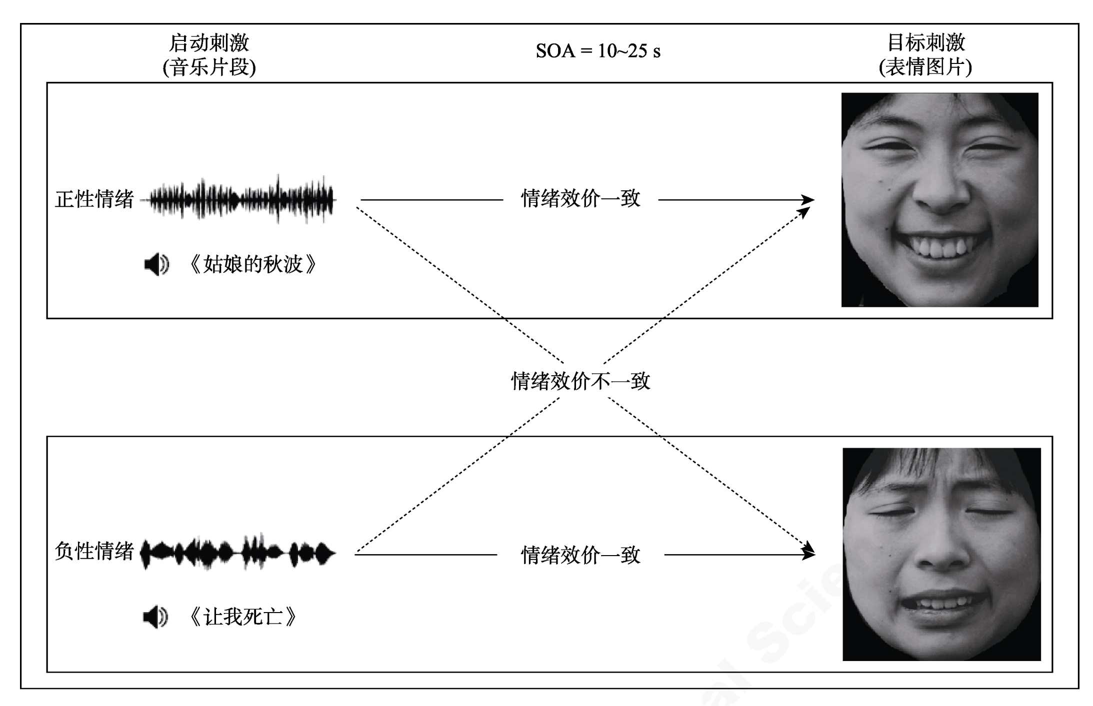


图 1　跨通道情绪启动范式例子

接受专业音乐训练 18 年以上)对声乐与器乐在弹性速度(M = 5.24, SD = 0.42)、力度(M = 5.38, SD = 0.40)、分句(M = 5.65, SD = 0.36)以及总体表演水平(M = 5.21, SD = 0.40)方面的评定具有较高的一致性(1 = 非常不一致, 7 = 非常一致), 其平均评定均高于 5。第三, 额外的 16 名普通大学生对音乐的情绪效价(1 = 非常不愉快, 4 = 中性, 7 = 非常愉快)和唤醒程度(1 = 非常平静, 4 = 中性, 7 = 非常激动)进行了评定。就情绪效价而言, 以情绪类型(正性情绪、负性情绪)和音乐类型(声乐、器乐)为被试内变量的方差分析显示, 情绪类型[$F(1, 29)$ = 2241.98, $p < 0.001$, η_p^2 = 0.99]与音乐类型[$F(1, 29)$ = 4.80, p = 0.037, η_p^2 = 0.14]的主效应以及二者的交互效应[$F(1, 29)$ = 8.37, p = 0.007, η_p^2 = 0.22]显著。简单效应分析表明, 就正性情绪音乐而言, 器乐(M = 5.43, SD = 0.30)与声乐(M = 5.46, SD = 0.26)的情绪效价没有差异, $F(1, 29)$ = 0.18, p = 0.674; 就负性情绪音乐而言, 器乐(M = 2.66, SD = 0.30)的情绪效价高于声乐(M = 2.41, SD = 0.35), $F(1, 29)$ = 13.40, p = 0.001, η_p^2 = 0.32。就唤醒程度而言, 情绪类型[$F(1, 29)$ = 111.87, $p < 0.001$, η_p^2 = 0.79]和音乐类型[$F(1, 29)$ = 419.34, $p < 0.001$, η_p^2 = 0.94]的主效应以及交互效应[$F(1, 29)$ = 18.97, $p <$ 0.001, η_p^2 = 0.40]显著。简单效应分析表明, 无论是正性情绪[$F(1, 29)$ = 36.07, $p < 0.001$, η_p^2 = 0.55]还是负性情绪[$F(1, 29)$ = 206.79, $p < 0.001$, η_p^2 = 0.88]音乐, 器乐的唤醒程度(正性情绪: M = 4.87, SD = 0.59; 负性情绪: M = 3.32, SD = 0.41)都低于声乐(正性情绪: M = 5.54, SD = 0.47; 负性情绪: M = 4.80, SD = 0.49)。

2.3　程序

首先, 电脑屏幕中心出现一个持续时间为 1 s 的黑色注视点“+”。注视点消失后, 启动刺激出现, 被试通过耳机聆听音乐。在音乐播放完毕之后, 呈现目标刺激 1 s。最后, 出现反应界面, 要求被试判断音乐与面孔图片所表达的情绪效价是否一致, 之后按空格键开始下一个试次。实验结束后, 要求被试报告音乐的标题, 但是他们未能报告出所有音乐标题中的任何一个关键词。

2.4　数据记录与分析

EEG 信号通过 NeuroScan Synamps 2 系统及其 64 导电极帽(10-20 方式排列电极)进行采集。左侧乳突为参考电极, 右侧乳突为记录电极。信号经过 AC 放大器放大, 滤波带通为 0.05~100 Hz, 采样率为 500 Hz。所有电极 R < 5 kΩ。

数据采用 Neuroscan 4.5 软件进行预处理。首

先，将原始数据转换为双侧乳突电极的平均参考，通过视觉观察手动排除肌电、心电、漂移等伪迹并使用软件的回归程序自动矫正眼电伪迹。其次，进行 0.1~30 Hz (24 dB/oct slope)的带通滤波。然后，对目标刺激−200~1000 ms 的脑电进行分段，并进行基线校正(−200~0 ms)。最后，自动剔除波幅在 ± 80 μV 之外以及反应错误的试次，对不同实验条件的试次分别进行叠加平均。

在预处理后，为了控制由众多电极和时间点进行多重比较引起的I类错误概率，我们使用 Fieldtrip 工具箱(Oostenveld, Fries, Maris, & Schoffelen, 2011)中的基于聚类的置换检验(cluster-based permutation test)方法(Maris & Oostenveld, 2007)进行统计分析。首先，使用相关样本 *t* 检验比较了情绪效价一致与不一致条件在每个相邻电极−时间点配对的差异。如果其差异在 0.05 水平上达到显著，那么该配对被作为一个簇(cluster)。对于每个簇，簇内每个电极−时间点 *t* 值之和作为簇水平的 *t* 值。然后，我们混合每个被试在情绪效价一致与不一致条件的数据，随机抽样 2000 次，计算出每次抽样得到的最大的簇水平统计量，从而构建一个假定 2 个实验条件没有差异的零分布。最后，对比实际观测到的与零分布得到的最大的簇水平统计量，使用蒙特卡罗方法估算 *p* 值($\alpha = 0.05$)。为了增加统计检验的敏感性，我们根据已有关于 N400 (Schirmer et al., 2005; Steinbeis & Koelsch, 2011)与 LPC (Herring et al., 2011; Zhang et al., 2010)的研究和目测，把面孔图片出现后的 200~500 ms 和 400~700 ms 分别作为 N400 和 LPC 的时间窗口。

3 结果

3.1 行为结果

我们以情绪效价一致性和音乐类型为被试内变量分别对正确率和反应时进行重复测量方差分析。就正确率而言，情绪效价一致性的主效应显著，$F(1, 16) = 13.50$, $p = 0.002$, $\eta_p^2 = 0.46$，表明听者在情绪效价一致配对上的成绩(M = 92.04%, SD = 5.93)高于在情绪效价不一致配对上的成绩(M = 81.27%, SD = 14.40)。同时，音乐类型的主效应也显著，$F(1, 16) = 17.15$, $p = 0.001$, $\eta_p^2 = 0.52$，表明听者在以声乐为启动刺激时对情绪效价一致性判断的成绩(M = 89.79%, SD = 8.94)比在以器乐为启动刺激时的成绩好(M = 83.52%, SD = 14.21)。然而，二者的交互作用不显著，$F(1, 16) = 0.02$, p = 0.903。

就反应时而言，情绪效价一致性的主效应显著，$F(1, 16) = 12.54$, $p = 0.003$, $\eta_p^2 = 0.44$，表明听者对情绪效价一致配对的反应(M = 993.40 ms, SD = 456.20)快于对情绪效价不一致配对的反应(M = 1174.97 ms, SD = 595.25)。但是，音乐类型的主效应[$F(1, 16) = 0.49$, $p = 0.495$]和二者的交互作用[$F(1, 16) = 0.27$, $p = 0.613$]均不显著。

3.2 ERP 结果

在 N400 时间窗口内，置换检验结果显示，当启动刺激是器乐时，情绪效价不一致的面孔图片比情绪效价一致的面孔图片诱发的脑电波更为负性(p = 0.011; 见图 2)。从图 3 可以看出，该效应主要产生于 281~471 ms 的右侧额颞和中央顶区电极。然而，当启动刺激是声乐时，情绪效价一致与不一致的面孔图片诱发的脑电波没有不同。

在 LPC 时间窗口内，置换检验结果表明，当启动刺激是声乐时，情绪效价不一致比一致的面孔图片诱发了一个波幅更大的正成分(p = 0.019; 见图 2)。从图 3 可以看出，该成分主要产生于 473~677 ms 的左侧额中和顶枕区电极。但是，当启动刺激是器乐时，情绪效价一致与不一致的面孔图片诱发的脑电波没有差异。

4 讨论

本研究使用 ERP 技术探讨声乐与器乐情绪加工的电生理反应是否存在差异。研究表明，与一致条件相比，与面孔情绪不一致的器乐曲诱发了 N400；而与面孔情绪不一致的声乐曲诱发了 LPC。这说明，声乐与器乐情绪加工的电生理反应存在差异。

我们发现，启动刺激为器乐片段时，情绪效价不一致比一致的面孔图片产生了更大的 N400。该结果与一些研究结果一致：无论启动刺激是单个和弦(Steinbeis & Koelsch, 2011; Zhou, Liu, Jiang, & Jiang, 2019)还是器乐片段(Goerlich et al., 2012; Zhang et al., 2019)，情绪不一致比一致的面孔图片或具有情绪意义的单词都会诱发大脑产生一个较大的(类似)N400。本研究的 N400 可能主要反映了大脑对情绪冲突信息的检测与整合(Zhang et al., 2006, 2010)。具体而言，启动刺激在概念水

图 2　启动刺激为器乐(a)与声乐(b)时面孔图片在 9 个代表性电极上诱发的平均脑电波形图
注：阴影区域表示每个时间点平均波幅的 95% CI。

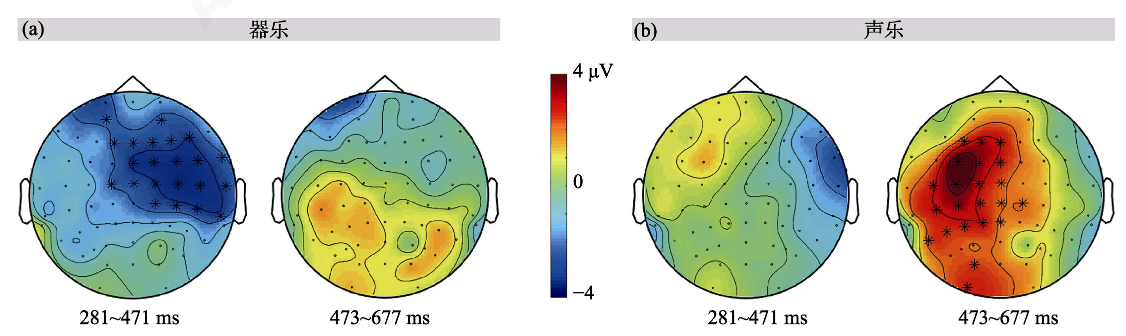


图 3　启动刺激为器乐(a)与声乐(b)时面孔图片在 2 个时间窗口产生的差异波的头皮分布地形图
注：*表示在特定时间窗口内 ERP 效应显著的电极。

平上预先激活了与目标刺激有关的情绪表征，从而使情绪一致的目标刺激诱发的 N400 波幅减小(Goerlich et al., 2012)。然而，对于情绪不一致的目标刺激，由于与其相关的情绪表征没有提前被激活，因此被试需要更多的认知资源对情绪信息进行整合(Zhang et al., 2010)，导致 N400 波幅增大。由于这种激活具有自动化的性质(Zhang et al., 2006)，因而 N400 与情绪信息的自动化整合加工有关。

另一方面，本研究的 N400 主要产生于右侧额颞和中央顶区电极，表明 N400 出现了右半球偏侧化现象。这与已有研究结果一致：当被试任务是判断目标刺激与启动刺激情绪(Schirmer et al., 2005)或意义(Painter & Koelsch, 2011)是否一致时，N400 出现右侧化。的确，fMRI 研究也表明，纯音乐比歌曲激活了更多的右侧脑区，比如额中回、额下回和扣带回(Brattico et al., 2011)。但是，本研究结果也与 ERP 研究结果矛盾：与启动和弦或器乐情绪不一致的目标刺激诱发的 N400 源于整个头皮(Steinbeis & Koelsch, 2011; Zhou et al., 2019)或前部区域(Goerlich et al., 2012)电极。这种矛盾可能源于被试任务的差异：在这些研究中，被试的任务是判断目标刺激的情绪效价；在本研究中，被试任务是判断目标刺激与启动刺激情绪效价的一致性，容易使被试较多关注启动刺激的情绪。未来研究需要对此进行探究。

我们也发现，启动刺激为声乐片段时，情绪效价不一致比一致的面孔图片诱发了更大的 LPC。该结果与先前研究结果一致：与启动歌曲(张伟霞 等, 2018)和言语韵律(Pell et al., 2015; 郑志伟，黄贤军, 2013)情绪不一致的面孔图片诱发了较大的 LPC。本研究的 LPC 可能反映了注意资源的分配(Zhang et al., 2010, 2012)。当启动刺激是歌曲时，由于歌曲本身的唤醒程度比较高，因而会引起被试较多的注意(Schupp et al., 2000)。同时，本研究的 LPC 也可能反映了情绪的整合加工(Herring et al., 2011)，因为听者需要先把歌词与音乐的情绪进行整合，然后再与面孔图片的情绪进行匹配。因为涉及言语信息，听者需要使用命题表征或思维有意识地对情绪信息进行加工(Strack & Deutsch, 2004)，所以体现了反思性的整合加工(Imbir, Spustek, & Żygierewicz, 2016)。然而，声乐情绪加工并没有产生 N400。该结果验证了张伟霞等人(2018)的结果：与歌曲情绪不一致的面孔图片没有诱发 N400。这可能是因为 N400 与 LPC 反映了不同的情绪整合功能。这种解释与情绪二元模型的观点一致(Jarymowicz & Imbir, 2015)。该模型认为，个体对刺激做出情绪反应的过程涉及 2 个独立的评价系统——自动评价系统和反思评价系统(Jarymowicz & Imbir, 2015)：前者不需要以言语为基础，后者则需要(Imbir et al., 2016)。

就 LPC 头皮分布而言，本研究的 LPC 主要产生于左侧额中和顶枕电极，表明了 LPC 的左半球偏侧化。这与已有研究结果一致：当单词情绪与反思评价系统相关(Imbir et al., 2016)或启动面孔与目标面孔情绪不一致(Cheal, Heisz, Walsh, Shedden, & Rutherford, 2014)时，左侧电极产生了较大的 LPC。fMRI 研究也表明，歌曲比纯音乐激活了更多的左侧脑区，比如颞上回、壳核、楔叶和小脑山坡(Brattico et al., 2011)。然而，本研究结果也与 ERP 研究结果不一致：与启动歌曲或言语韵律情绪不一致的面孔图片诱发的 LPC 产生于整个头皮(张伟霞 等, 2018)或双侧顶枕区(Pell et al., 2015; 郑志伟，黄贤军, 2013)电极。这种矛盾可能与不同研究使用的启动刺激材料情绪效价和唤醒程度的差异有关。由于这些研究没有提供相关数据，我们难以确定产生矛盾结果的具体原因。

致谢：感谢上海师范大学音乐学院蒋存梅教授在本文写作中提供的帮助；感谢上海公共卫生临床中心蔡丹超博士、中国科学院心理研究所孙丽君博士与重庆邮电大学李沛洋博士在 ERP 数据分析和作图方面提供的帮助。

参考文献

龚栩，黄宇霞，王妍，罗跃嘉. (2011). 中国面孔表情图片系统的修订. *中国心理卫生杂志, 25*(1), 40–46.

张伟霞，王莞琪，周临舒，蒋存梅. (2018). 歌词对音乐情绪加工的影响：行为与 ERP 研究. *心理学报, 50*(12), 1346–1355.

郑志伟，黄贤军. (2013). 情绪语音调节面孔表情的识别：ERP 证据. *心理科学, 36*(1), 33–37.

Ali, S. O., & Peynircioğlu, Z. F. (2006). Songs and emotions: Are lyrics and melodies equal partners? *Psychology of Music, 34*(4), 511–534.

Brattico, E., Alluri, V., Bogert, B., Jacobsen, T., Vartiainen, N., Nieminen, S., & Tervaniemi, M. (2011). A functional

MRI study of happy and sad emotions in music with and without lyrics. *Frontiers in Psychology, 2*, 308. doi: 10.3389/fpsyg.2011.00308.

Cheal, J. L., Heisz, J. J., Walsh, J. A., Shedden, J. M., & Rutherford, M. D. (2014). Afterimage induced neural activity during emotional face perception. *Brain Research, 1549*, 11–21.

Diamond, E., & Zhang, Y. (2016). Cortical processing of phonetic and emotional information in speech: A cross-modal priming study. *Neuropsychologia*, *82*, 110–122.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Föcker, J., & Röder, B. (2019). Event-related potentials reveal evidence for late integration of emotional prosody and facial expression in dynamic stimuli: An ERP study. *Multisensory Research*, *32*(6), 473–497.

Franco, F., Chew, M., & Swaine, J. S. (2017). Preschoolers' attribution of affect to music: A comparison between vocal and instrumental performance. *Psychology of Music*, *45*(1), 131–149.

Goerlich, K. S., Witteman, J., Schiller, N. O., van Heuven, V. J., Aleman, A., & Martens, S. (2012). The nature of affective priming in music and speech. *Journal of Cognitive Neuroscience, 24*(8), 1725–1741.

Hajcak, G., MacNamara, A., & Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: An integrative review. *Developmental Neuropsychology*, *35*(2), 129–155.

Hajcak, G., & Olvet, D. M. (2008). The persistence of attention to emotion: Brain potentials during and after picture presentation. *Emotion*, *8*(2), 250–255.

Herring, D. R., Taylor, J. H., White, K. R., & Crites Jr, S. L. (2011). Electrophysiological responses to evaluative priming: The LPP is sensitive to incongruity. *Emotion, 11*(4), 794–806.

Imbir, K. K., Spustek, T., & Żygierewicz, J. (2016). Effects of valence and origin of emotions in word processing evidenced by event related potential correlates in a lexical decision task. *Frontiers in Psychology, 7*, 271. doi: 10.3389/fpsyg.2016.00271.

Jarymowicz, M. T., & Imbir, K. K. (2015). Toward a human emotions taxonomy (based on their automatic vs. reflective origin). *Emotion Review, 7*(2), 183–188.

Kraus, D., & Horowitz-Kraus, T. (2014). The effect of learning on feedback-related potentials in adolescents with dyslexia: An EEG-ERP study. *PLoS ONE, 9*(6), e100486. doi: 10.1371/journal.pone.0100486.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*(1), 621–647.

Loui, P., Bachorik, J. P., Li, H. C., & Schlaug, G. (2013). Effects of voice on emotional arousal. *Frontiers in Psychology, 4*, 675. doi: 10.3389/fpsyg.2013.00675.

Luck, S. J. (2014). *An introduction to the event-related potential technique* (2nd ed.). Cambridge, MA: The MIT Press.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164*(1), 177–190.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience,* 1-9. doi: 10.1155/2011/156869.

Painter, J. G., & Koelsch, S. (2011). Can out-of-context musical sounds convey meaning? An ERP study on the processing of meaning in music. *Psychophysiology, 48*(5), 645–655.

Paulmann, S., & Pell, M. D. (2010). Contextual influences of emotional speech prosody on face processing: How much is enough? *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 230–242.

Paulmann, S., & Uskul, A. K. (2017). Early and late brain signatures of emotional prosody among individuals with high versus low power. *Psychophysiology*, *54*(4), 555–565.

Pell, M. D., Rothermich, K., Liu, P., Paulmann, S., Sethi, S., & Rigoulot, S. (2015). Preferential decoding of emotion from human non-linguistic vocalizations versus speech prosody. *Biological Psychology*, *111*, 14–25.

Schirmer, A., Kotz, S. A., & Friederici, A. D. (2005). On the role of attention for the processing of emotions in speech: Sex differences revisited. *Cognitive Brain Research*, *24*(3), 442–452.

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology*, *37*(2), 257–261.

Steinbeis, N., & Koelsch, S. (2011). Affective priming effects of musical sounds on the processing of word meaning. *Journal of Cognitive Neuroscience, 23*(3), 604–621.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*(3), 220–247.

Stratton, V. N., & Zalanowski, A. H. (1994). Affective impact of music vs. lyrics. *Empirical Studies of the Arts, 12*(2), 173–184.

Zhang, Q., Kong, L., & Jiang, Y. (2012). The interaction of arousal and valence in affective priming: Behavioral and electrophysiological evidence. *Brain Research, 1474*, 60–72.

Zhang, Q., Lawson, A., Guo, C., & Jiang, Y. (2006).

Electrophysiological correlates of visual affective priming. *Brain Research Bulletin*, *71*(1-3), 316–323.
Zhang, Q., Li, X., Gold, B. T., & Jiang, Y. (2010). Neural correlates of cross-domain affective priming. *Brain Research*, *1329*, 142–151.
Zhang, W. X., Liu, F., Zhou, L. S., Wang, W. Q., Jiang, H. Y., & Jiang, C. M. (2019). The effects of timbre on neural responses to musical emotion. *Music Perception, 37*(2), 134–146.
Zhou, L. S., Liu, F., Jiang, J., & Jiang, C. M. (2019). Impaired emotional processing of chords in congenital amusia: Electrophysiological and behavioral evidence. *Brain and Cognition, 135,* 103577. doi: 10.1016/j.bandc.2019.06.001.

Emotional processing in vocal and instrumental music: An ERP study

JIANG Jun[1]; ZHANG Weixia[2]; WANG Wanqi[3]
([1] *Music College, Shanghai Normal University, Shanghai* 200234, *China*)
([2] *College of Education, Shanghai Normal University, Shanghai* 200234, *China*)
([3] *School of Finance and Business, Shanghai Normal University, Shanghai* 200234, *China*)

Abstract: The event-related potential (ERP) technique was used to investigate whether there are different neural responses to musical emotion when the same melodies are presented in the voice and instrumental timbre such as the violin. With a crossmodal affective priming paradigm, target faces were primed by affectively congruent or incongruent vocal and instrumental music. Participants were asked to judge whether the prime-target pair was affectively congruent or incongruent. The results revealed a larger late positive component (LPC) at the time window of 473~677 ms in response to affectively incongruent versus congruent trials in the vocal version, whereas a larger N400 effect at the time window of 281~471 ms was observed in the instrumental version. These results indicate differential patterns of neurophysiological responses to emotion processing of vocal and instrumental music.

Key words: vocal music; instrumental music; emotion processing; N400; LPC